

IMBE and AMBE Speech Compression

Robert Rudolph
Digital Voice Systems, Inc.

Eddie Yu
Gold Insignia Electronic Co., Ltd.

The Improved Multi-Band Excitation (IMBE™) and the Advanced Multi-Band Excitation (AMBE™) speech compression algorithms are the performance leaders for low-bit-rate speech compression systems. MBE-based technology provides superior speech quality, while requiring substantially fewer MIPS and less memory than other speech coders. In addition, the speech coders have been designed for robustness in both background noise and channel errors. The IMBE™ and AMBE® speech coders have been selected for many international mobile communication standards, including standards in satellite communications, commercial aircraft telephony, and digital mobile radio. They are also widely used in many other applications such as in secure communications, voice storage, and desktop video conferencing.

DVSI's IMBE™ and AMBE® speech coders are undefeated in numerous independent evaluations, winning 6 out of the last 6 evaluations. These systems have been consistently shown to provide the highest performance of any speech coder available. For example, in 1992 four real-time 7.2 kbits/sec speech coders were

evaluated by the Telecommunications Industry Association (TIA) for the purpose of selecting a speech coder for the APCO Project 25 North American land mobile radio communication system. DVSI's 7.2 kbits/sec IMBE™ speech coder contained 4.4 kbits/sec of speech coding, and 2.8 kbits/sec of error correction coding. In addition to DVSI's IMBE™ speech coder, the participants included VSELP, STC, and CELP. The results of this test are shown in Figure 1.

In addition to showing significantly better voice quality in nearly every test condition, the IMBE™ speech coder demonstrated other significant advantages, such as computational simplicity. Based on these impressive test results, the IMBE™ speech coder was selected by the TIA as the voice coding standard for APCO Project 25.

Based on IMBE™'s high speech quality and robustness to errors, it has become the international standard for a number of communication systems. The IMBE™ system is currently the standard for sev-

eral global satellite-based mobile communication services, including several Inmarsat and OPTUS services. The performance of DVSI's technology in the field has been proven and has even exceeded expectations. The IMBE™ speech coder has successfully been in use commercially since the late 1980's.

DVSI continues to improve its lead in speech compression technology, and has recently introduced the Advanced Multi-Band Excitation (AMBE™) speech coder. The AMBE™ technology builds on the IMBE™ speech coder and provides further improvements in speech quality and robustness. In 1994, the AMBE™ speech coder was independently tested by Inmarsat, and the results showed that the AMBE® speech coder had significant advantages over other technologies. In fact, the performance of 3.6 kbits/sec AMBE® system was compared to the performance of the full-rate (8 kbits/sec) VSELP North American digital cellular standard (IS-54) (where the bit rates refer to the speech

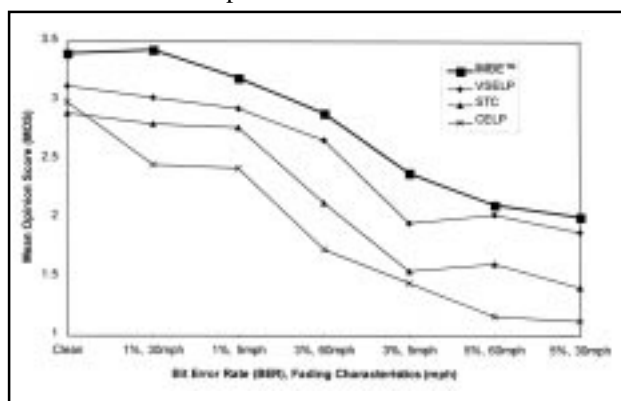


Figure 1. APCO Voice Quality Test - 1992

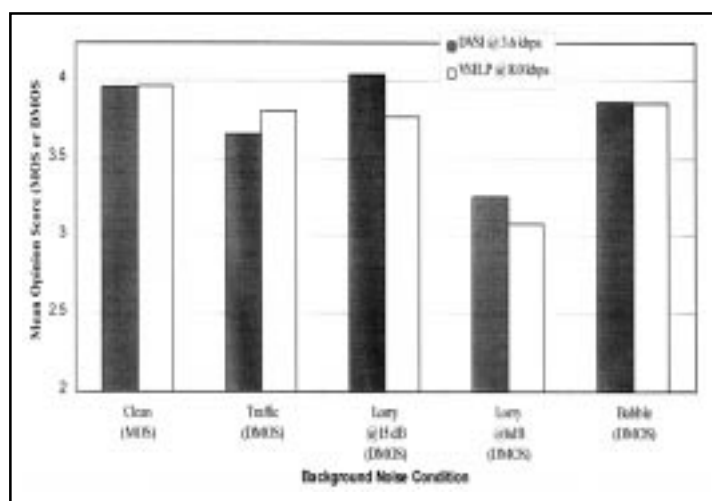


Figure 2. Inmarsat Voice Quality Test - 1994

coders without error correction). The 3.6 kbits/sec AMBE[®] system was shown to have better overall performance, especially in background noise, than the VSELP coder operating at over twice the data-rate, as shown in figure 2.

The MBE technology

One of the major factors that has contributed to the success of the IMBE[™] and AMBE[®] speech coders is that they use a fundamentally different technology than standard speech coders. The technology is the outgrowth of work begun at the Massachusetts Institute of Technology in the early 1980's. The goal of this work was to develop a robust speech model which would outperform the linear prediction speech model used in traditional speech coders. The result of this work was the Multi-Band Excitation (MBE) speech model. This speech model provides a unique speech coding framework which results in a number of advantages over linear prediction based speech coders such as CELP, RELP, VSELP, LPC-10, etc...

Most CELP speech coders make a single determination as to whether each speech segment is a periodic (voiced) signal, or a noise-like (unvoiced) signal. One major difference between CELP coders and the MBE speech coder is that the MBE coder divides each segment of speech into distinct frequency bands and makes a voiced/unvoiced (V/UV) decision for each frequency band. This allows the excitation signal for a particular speech segment to be a mixture of periodic (voiced) and noise-like (unvoiced) energy. This added degree of freedom in the modelling of the excitation signal allows the MBE speech model to generate higher quality speech than conventional speech models. In addition, it allows the MBE speech model to be robust in the presence of background noise.

The inherent problem with linear prediction based speech coders is that the linear prediction model does not yield high quality speech (or robustness to background noise) without the addition of a prediction residual. The prediction residual can be viewed as an error signal which corrects for inaccuracies in the linear prediction model. Elimination of this residual, as is done in the government standard 2.4 kbits/sec LPC-10 system, results in a harsh, mechanical quality in the speech. Consequently, all high quality linear predictive speech coders transmit a residual. The primary difference between these systems is the manner in which they accomplish this task. The favored method used in linear predictive speech coding at rates below 8 kbits/sec

is to divide the residual into small pieces or vectors and to then search through a codebook to find the code vector which is the closest match. Unfortunately, searching through a reasonable sized codebook is a computationally complex task. Furthermore, a particular codebook is designed to operate at a fixed data rate and is not easily scalable to other data rates.

The IMBE[™] and AMBE[®] speech coders do not have these problems because they are not based upon linear prediction. Instead, they use the Multi-Band Excitation speech model to produce high quality speech without the need for a residual signal. They maintain speech intelligibility and naturalness at rates as low as 2.4 kbits/sec. In addition, the MBE[™] speech coders do not require the use of codebooks. Consequently, the IMBE[™] system requires fewer computations than either CELP or VSELP. Finally, the IMBE[™] and AMBE[®] speech coders can be easily scaled to virtually any data rate above 2.4 kbits/sec.

Products and applications

A full-line of IMBE[™] and AMBE[®] software and hardware products have been produced by DVSI. Efficient software implementations are currently available on many of the most popular fixed-point and floating-point Digital Signal Processor (DSP) families. DVSI offers real-time software which converts one single, low-cost signal processing chip into a full-duplex IMBE[™] voice codec. Because of the low complexity of the DVSI speech coders, more efficient and therefore lower-cost implementations can be produced.

Low-cost, full-duplex voice codec modules and chips have also been produced by DVSI and are currently in use commercially throughout the world. This year, DVSI is introducing the first single chip vocoder offering high speech quality at data-rates as low as 2.4 kbits/sec. These AMBE[®] and IMBE[™] hardware products are ideal for a wide variety of applications. A number of advanced features including error correction, voice/silence detection, DTMF detection/synthesis, echo cancellation, and soft-decision decoding have been incorporated into these products.

The IMBE[™] speech coder is in use today in many real-world applications. One example is in the area of mobile satellite phones. These portable satellite phones are designed to provide worldwide mobile phone service from almost anywhere in the world, whether on land, at sea, or in the air. These phones transmit and receive directly from a satellite,

therefore, they are especially useful in areas that do not have any cellular phone services. Example users include: journalists, reporters, business travelers, emergency service personnel, and disaster relief personnel. The phones may be used as portable or fixed units, or they may be used to provide mobile communication services in automobiles, trucks, ships (yachts, tankers, cargo ships, cruise ships), and aircraft.

Another major application area of IMBE[™] is in digital land mobile radios. The IMBE[™] speech coder is the new digital standard for public safety radio in North America (APCO Project 25). Many organizations are currently upgrading their analog radios to higher performance digital radios. The IMBE[™] speech coder and error correction coding is an integral part of many new digital radio systems. In today's congested spectral environment, the IMBE speech coder offers the ability to improve performance while reducing spectral utilization by a factor of 4. These radios are used by organizations such as police departments, fire departments, emergency service personnel, trucking fleets, delivery fleets, and taxi cab fleets. Land mobile radios include both vehicle-based and hand-held ("walkie-talkie") units.

The IMBE[™] speech coder has been used with great success in desktop videoconferencing applications. The high compression ratios of the speech coder allow data and/or video communications to occur simultaneously with speech communications. Because of the limited bandwidth available over standard analog phone lines, compression technologies are of prime importance in these types of applications. With IMBE[™] speech compression, high quality voice is maintained at the same time that users are file-sharing, working on a "whiteboard", and sending video, all at the same time over a single analog phone line.

New digital answering machines can benefit from many of the advantages of DVSI's speech coder. The technology allows better speaker recognition as well as better speech quality at very low data rates. Users are able to recognize the caller simply by the sound of their voice. The variable rate capabilities of the coder can be used to allow customers to vary the storage capacity of their machine as needed. The number of messages that may be stored and the amount of time available per message can be increased significantly by using the 2400bps AMBE[™] speech coder. No additional RAM is required to store many times more messages than previously possible.

In summary, the MBE speech coders

use a revolutionary new speech model to achieve better performance than other speech coders. Their primary advantage is that superior speech quality is achieved at low data-rates. They are also less complex and therefore can be implemented very cost-effectively. The IMBE™ and AMBE® speech coders have been proven in the field in many applications and they have been chosen as the standard for several major international communication systems.

References

- [1] D. W. Griffin and J. S. Lim, "A Speech Spectral Analysis/Synthesis System," Proc. of 1984 DSP Workshop, pp. 5.1.1-5.1.2, October 1984.
- [2] D. W. Griffin and J. S. Lim, "A New Model-Based Speech Analysis/Synthesis System," Proc. of ICASSP 1985, Tampa, FL., pp. 513-516, March 1985.
- [3] D. W. Griffin and J. S. Lim, "Multi-band Excitation Vocoder," IEEE Transactions on ASSP, Vol. 36, No. 8, August 1988.
- [4] J. C. Hardwick and J. S. Lim, "A 4.8 kbits/sec Multi-Band Excitation Speech Coder," Proc. ICASSP 88, New York, N.Y., pp. 374-377, April 1988.
- [5] J. C. Hardwick and J. S. Lim, "A 4800 bps Improved Multi-Band Excitation Speech Coder," Proc. of IEEE Workshop on Speech Coding for Telecommunications, Vancouver, Canada, Sept. 5-8, 1989.
- [6] J. C. Hardwick and J. S. Lim, "The Application of the IMBE Speech Coder to Mobile Communications," Proc. ICASSP 91, Toronto, Canada, pp. 249-252, May 1991.
- [7] S. W. Wong, "An Evaluation of 6.4 kbit/s Speech Codecs for INMARSAT-M System," Proc. ICASSP 91, Toronto, Canada, pp. 629-632, May 1991.